

METHOD OF FORMING FLASH MEMORIES WITH HIGH COUPLING
RATIO AND THE STRUCTURE OF THE SAME

The present invention relates to a semiconductor
5 device, and more specifically, to a method of
fabricating flash memories and the structure of the
same.

Background of the Invention

10 Various nonvolatile memories have been disclosed
in the prior art. For example, Mitchellx has proposed
EPROMs with self-aligned planar array cell. In this
technique, buried diffusion self-aligned to the
floating gate avalanche injection MOS transistors are
15 used for the bit lines. Cross point array technology
has been disclosed. The self-aligned source and drain
will allow this device to be optimized even further
for programming speed. See A. T. Mitchellx, "A New
Self-Aligned Planar Cell for Ultra High Density
20 EPROMs", IEDM, Tech. pp. 548-553, 1987".

Flash memory is one of the segments of
nonvolatile memory devices. The device includes a
floating gate to storage charges and an element for
25 electrically placing charge on and removing the
charges from the floating gate. One of the
applications of flash memory is BIOS for computer.
Typically, the high-density nonvolatile memories can
be applied as the mass storage of portable handy
30 terminals, solid state camera and PC cards. That is
because that the nonvolatile memories exhibit many
advantages, such as a fast access time, low power
dissipation, and robustness. Bergemont proposed

another cell array for portable computing and telecommunications application, which can be seen in Bergmont et al., "Low Voltage NVGTM: A New High Performance 3 V/5 V Flash Technology for Portable
5 Computing and Telecommunications Applications", IEEE Trans. Electron Devices, vol. ED-43, p. 1510, 1996. This cell structure is introduced for low voltage NOR Virtual Ground (NVG) flash memory having fast access time. In the flash array schematic, field oxides (FOX)
10 are formed between cells such that a poly extension on FOX of each cell provides adequate gate coupling ratio. Bergmont also mentioned that the portable telecommunications and computing have become a major driving force in the field of integrated circuits. In
15 the article, the access time is one of the key concerns for low voltage read operation. The NVG array uses select devices to achieve a fast access time by reducing the pre-charge time to that of a single segment rather than the full bit-line.

20

The formation of nonvolatile memories toward the trends of low supply power and fast access, because these requirements are necessary for the application of the mobile computing system. Flash memory needs
25 the charges to be hold in the floating gate for a long periods of time. Therefore, the dielectric that is used for insulating the floating gate needs to be high performance. At present, the low voltage flash memory is applied with a voltage of about 3V or 5V
30 during charging or discharging the floating gate. As known in the art, tunneling is a basic technology in charging or discharging. In order to attain high tunneling efficiency, the thickness of the dielectric

between the floating gate and substrate have to be scaled down due to the supply voltage is reduced.

United States Patent No. 6,180,459 to Sheu,
5 entitled "Method for fabricating a flash memory with shallow trench isolation", filed on January 8, 1999. The prior art disclosed a method for fabricating a flash memory comprising forming a shallow trench isolation (STI) structure is also formed in the
10 method. A further United States Patent No. 6,172,395 to Chen, et al., entitled "Method of manufacture of self-aligned floating gate, flash memory cell and device manufactured thereby" and assigned to Taiwan Semiconductor Manufacturing Company (Hsin-Chu, TW).

15
A further prior article can be seen in United States Patent No. 6,171,909 to Ding, et al., entitled "Method for forming a stacked gate" and assigned to United Semiconductor Corp. (Hsinchu, TW) and the
20 prior art is filed on April 16, 1999. The method includes forming a first dielectric layer, a conductive layer and a silicon nitride layer sequentially over a substrate. A photoresist pattern is formed over the silicon nitride layer. The silicon
25 nitride layer, conductive layer, first dielectric layer and substrate are etched by using the photoresist pattern as an etching mask until forming a plurality of trenches in the substrate. An insulating layer is formed over the substrate. A
30 conductive spacer is formed on the sidewalls of the conductive layer and silicon nitride layer, wherein the conductive spacer and conductive layer serve as a first gate conductive layer. The silicon nitride

layer is removed. A second dielectric layer and a second gate conductive layer are formed over the substrate. The second gate conductive layer, second dielectric layer and first gate conductive layer are
5 patterned to form a control gate, a patterned dielectric layer and a floating gate, respectively.

Summary of the Invention

10 The object of the present invention is to form a self-aligned floating gate with higher coupling ratio.

The further object is to form the floating gate and trench without switching the chamber.

The yet object of the present invention is to
15 form the floating gate and trench, simultaneously.

The present invention comprises forming a tunneling dielectric layer on the substrate and a first conductive layer on the tunneling dielectric
20 layer. Then, patterning the first polysilicon layer, the tunneling dielectric layer and the substrate to form trenches therein before a gap-filling material is refilled into the trenches and over the substrate. A portion of the gap-filling material is removed to
25 form trench isolations. A portion of the trench isolations is etched to form slots between the etched first conductive layer. Next, a second conductive layer is formed over a surface of the slots and the etched first conductive layer. The second conductive
30 layer is etched, thereby forming sidewall spacers on the slot. Next step is to form a second dielectric layer on the trench isolation, the sidewall spacers and the first conductive layer; and forming a third

conductive layer on the second dielectric layer to act as a control gate.

5 The flash memory structure includes a substrate having trenches formed therein, a first dielectric layer and a first conductive layer are stacked on the substrate. Isolations are formed in the trenches and protruding over the surface of the substrate, wherein the first conductive layer is also protruded over the
10 isolations. A second conductive layer is lying the surface of the first conductive layer and a second dielectric layer formed thereon. A third conductive layer is formed on the second dielectric layer. The floating gate is consisted of first conductive layer
15 and the second conductive layer.

Brief Description of the Drawings

20 The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same becomes better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:

25 FIGURE 1 is a cross sectional view of a semiconductor wafer illustrating the steps of forming trenches in a semiconductor substrate according to the present invention.

30 FIGURE 2 is a cross sectional view of a semiconductor wafer illustrating the step of forming trench isolation according to the present invention.

FIGURE 3 is a cross sectional view of a semiconductor wafer illustrating the steps of etching

the isolations and forming floating gate according to the present invention.

FIGURE 3A is a cross sectional view of a semiconductor wafer illustrating the FIGURE 3B is a cross sectional view of a semiconductor wafer illustrating the step of forming sidewall spacers according to the present invention.

FIGURE 4-4B are cross sectional views of a semiconductor wafer illustrating the step of forming dielectric layer and control gate according to FIGURE 3-3A of the present invention.

FIGURE 5 is a further cross sectional view of the present invention.

Detailed Description of the Preferred Embodiment

The present invention proposes a novel method to fabricate a flash nonvolatile memory. In the method, trench and the floating gate are simultaneously formed in one step. The structure can form self-aligned silicon island structure. During the formation of trench, the substrate does not need to be switched to another chamber due to the floating gate and substrate are composed of identical material. The detail description will be seen as follows.

The first embodiment of the structure, as shown in FIGURE 4, includes a substrate 2 having trenches formed therein, a first dielectric layer 4 and a first conductive layer are stacked on the substrate 2. Isolations 10 are formed in the trenches and protruding over the surface of the substrate 2, wherein the first conductive layer 6 is also protruded over the isolations 10. A second conductive

layer 14 is lying the surface of the first conductive layer 6 and a second dielectric layer 18 formed thereon. A third conductive layer 20 is formed on the second dielectric layer 18. The floating gate is
5 consisted of first conductive layer 6 and the second conductive layer 18.

Alternatively, further sidewall spacers 16 are attached on the side wall of the floating gate
10 consisting of first conductive layer 6 and the second dielectric layer 18, as shown in FIGURE 4A.

Another embodiment can be seen in FIGURE 4b, the second conductive layer 14 is omitted and the spacers
15 14a are attached on the floating gate consisting of first conductive layer 6 only. Yet another embodiment of the present invention is that even the spacers 14a are also abridged, referring to 5.

20 The processes for forming the devices are described as follows.

A semiconductor substrate is provided for the present invention. In a preferred embodiment, as
25 shown in the FIGURE 1, a single crystal silicon substrate 2 with a $\langle 100 \rangle$ or $\langle 111 \rangle$ crystallographic orientation is provided. A thin tunneling dielectric layer 4 consisted of silicon oxide is formed on the substrate 2. Typically, the oxide 4 can be grown in
30 oxygen ambient at a temperature of about 700 to 1100 degrees centigrade. Other method, such as chemical vapor deposition, can also form the gate oxide 4. In the embodiment, the thickness of the silicon dioxide

layer 4 is approximately 15-250 angstroms. Subsequently, a conductive layer such as doped polysilicon layer 6 or metal is deposited on the tunneling dielectric layer 4. Generally, the
5 polysilicon layer 6 is chosen from doped polysilicon or in-situ polysilicon. For an embodiment, the doped polysilicon layer 8 is doped by phosphorus using a PH_3 source.

10 Next, still referring to FIGURE 1, standard lithography and etching steps are used to etch the polysilicon layer 6, tunneling dielectric layer 4 and into the substrate 2 to a depth to form trenches 8 in the substrate. In the step, the isolation area and
15 the floating gate 6 are simultaneously formed. This method provides self-aligned floating gate. In this case, the dry etching uses $\text{CF}_4 + \text{O}_2$ plasma. A gap filling material 10, such as oxide formed by a CVD system, which is referred to as CVD-oxide, is
20 refilled into the trenches 8 for isolation. Preferably, the oxide layer 10 is formed at a temperature in the range of about 400 to 600 degrees centigrade. A chemical mechanical polishing (CMP) technology is used to remove the top of the CVD-oxide
25 10 to the surface of the conductive layer 6.

Referring to FIGURE 3, an etching back process is performed to etch the oxide 10 using the high selectivity etching between oxide and silicon. The
30 upper portion of the oxide 10 is removed to form slots 12 between the floating gates 10, and the depth of the slots 12 from the surface of the floating gate 10 can be defined by the etching. A second conductive

layer 14 is conformally formed along the surface of the floating gate 6 and the slot 12. The conductive layer 6 and the second conductive layer 14 are combined to act as the floating gate. If the width of the first conductive 6 is x , the thickness of the second conductive layer 14 is Y , thus the upper surface area of the floating gate is approximately $X+2Y$. It can improve the coupling ratio. An etching is optionally used to etch the second conductive layer 14 to form spacers 14a on the side wall of the slot 12, as shown in FIGURE 3B.

Another embodiment is shown in FIGURE 3A, sidewall spacers 16 are formed in the slot 12 and attached on the sidewall of the second conductive layer 14. The sidewall spacers 16 are formed by well known anisotropically etching. Assume again that the thickness of the sidewall spacers 16 is Z , then the upper surface of the floating gate is about $X+2Y+2Z$.

Next, please turn to FIGURE 4-4B, an inter polysilicon dielectric (IPD) 18 is formed at the top of the floating gate. Preferably, the ONO (oxide/nitride/oxide) or NO is used as the IPD 18. Finally, a third conductive layer 20, such as doped polysilicon layer or metal 16, is formed on the IPD 18 as control gate. The doped polysilicon layer 20 can be chosen from doped polysilicon or in-situ doped polysilicon. In addition, the metal or alloy layer can be used as the conductive layer 20.

Turning to FIGURE 5, alternatively, the dielectric layer may directly be formed 18 on the

surface of the etched first conductive layer 6, followed by forming the layer 20.

5 In the process of this invention, a tunnel oxide layer, floating gate and the trench are defined at the same time. Namely, the floating gate and tunnel oxide simultaneously are self-aligned formed adjacent to the shallow trench isolation. Further, the present invention provides a benefit of that the substrate
10 does not need to be switched to another chamber during the trench etching due to the floating gate and substrate are composed of the same material.

15 As will be understood by persons skilled in the art, the foregoing preferred embodiment of the present invention is illustrative of the present invention rather than limiting the present invention. Having described the invention in connection with a preferred embodiment, modification will now suggest
20 itself to those skilled in the art. Thus, the invention is not to be limited to this embodiment, but rather the invention is intended to cover various modifications and similar arrangements included within the spirit and scope of the appended claims, the scope of which should be accorded the broadest
25 interpretation so as to encompass all such modifications and similar structures.

30 While the preferred embodiment of the invention has been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.